



## Comment on **essd-2021-213**

Anonymous Referee #3

---

Referee comment on "Two decades of flask observations of atmospheric  $\delta\text{O}_2/\text{N}_2$ ,  $\text{CO}_2$ , and APO at stations Lutjewad (the Netherlands) and Mace Head (Ireland) plus 3 years from Halley station (Antarctica)" by Linh N. T. Nguyen et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-213-RC3>, 2021

---

In this manuscript Nguyen et al. present ~20 years of observations of atmospheric oxygen and carbon dioxide from two Northern Hemisphere flask sampling stations and 3 years of observations of the same from an Antarctic station. The authors are presenting a substantial dataset of value to the greater O<sub>2</sub> and CO<sub>2</sub> communities. The data are openly accessible from the ICOS Carbon Portal, and the files are self-explanatory (with one minor exception). The manuscript is of good quality and I recommend publication after in ESSD after addressing some of the minor points I raise below.

### MAJOR COMMENTS

1. Missing details: Some of the important details pertaining to the measurements are not presented. The reader is instead pointed to some relevant citations. This forces the data user to read this paper, plus 2-3 more, including a PhD thesis. I think reproducing some of the key details here would be a service to the reader, particularly details on the flask design/shape, and more information about the mass spec analysis. I am also surprised there are no relevant changes to mention during 20 years of sampling in either sample collection, analytical approach, gas handling, storage, etc, or changes to the automatic flask sampler. Also useful would be details on the different tanks used. Maybe I missed this, but I did not see anything about the valve type and seal, volume, interior, etc. Could the authors include some kind of change log, or table of notable events? If there is truly nothing to mention, I applaud the authors' consistency over 20 years of sampling!

2. Uncertainty and data quality: There is not much in the way of uncertainty analysis or constraint. I would say this is the biggest shortcoming of this paper. The reader is left with no real guidance as to how to assess the uncertainty in the individual measurements, or, perhaps more importantly, the trend. It does not seem like they have a good handle on the uncertainty due to primary tank drift (e.g. Keeling et al 2006 Tellus 59). A full uncertainty analysis may be out of scope, but they could at least put some constraints with the data at hand on the long-term trend and the reproducibility of a given flask measurement. As I see it, if someone wanted to use these flask records in some kind of analysis, this paper would be the main source of guidance. As such, I hope the authors can provide a bit more help in how an interested user could constrain the uncertainty of the measurements.

3. Supporting data: The authors are publishing sample time, sample height, and analyte concentration, but further data on the flasks is not included: analysis date, fill pressure, average flow rate, temperature data for sampling and analysis if it exists, etc. This supporting data would be helpful for anyone interested in further QA/QCing or using the data. The authors also do not include the other species used to filter/select the samples for background conditions. I suggest the authors should also seriously consider supplying the non-background data with flags, instead of only the background samples, and the CO/Radon data used to filter them. Or, at least provide DOIs as to where one could find it.

#### MINOR COMMENTS

Data files: It is not clear to me from the paper or the header what exactly the standard deviation column represents in the data files. I suggest the authors add this to the header, or put it in a subsection in the paper describing the files.

L17-18: better to provide a metric here than to use the subjective "high-quality". Also, I am not sure if inter-comparisons tell us anything about the quality of the calibration--all of the labs could be making the same mistake.

L19: suggest striking the "internationally-recognised" for the sake of brevity.

L25: Compatible can only be assessed if two measurements are made on the same air (tank or background), so I don't think it is correct to say that seasonal cycles are compatible if measured at different locations. Better to say they are in good agreement.

L40 - "a strong aide" -- Curious wording

L50-63: suggest cutting this entire paragraph up to "Our Laboratory...", and combining with the next paragraph.

L81: What does the "(formerly)" mean? It used to be called this but the name has changed, or it's not operational now? Please clarify.

L94: Nafion driers are not very common in O<sub>2</sub>/N<sub>2</sub> measurements. Does Nafion fractionate O<sub>2</sub>/N<sub>2</sub>? If the authors have tested this, I would encourage them to include such results here (or provide a citation).

L115: Please provide a plumbing diagram(s) of the flask samplers.

L116: Could you include a drawing or picture of one of the flasks? Do they have dip tubes?

L123: Dried to what dewpoint? Please include specifications on the cryotrap(s).

L124: What is the flow rate during flask sampling?

L124: Atmospheric pressure varies, please give exact fill pressure with observed range.

L131: Please give full details on sampling protocols for Mace Head.

L170: "Relatively very stable" -- ambiguously worded

L191: The influence of fossil fuel burning on APO is not small--that is why there is a large

trend in APO.

L210: Is there a systematic difference between first, second, and third analyses? Why sometimes 2 and sometimes 3?

L215: I don't fully follow -- you are assigning the WT a value and then assigning flasks a value based on comparison with the flask? Or flasks are assigned values from the MREF and then corrected for long-term drift through the WT?

L249: From Figure 3 it looks like some of the drift is not well-described by an average drift rate. Can the authors comment on this?

L264: I am still a little confused about how values are assigned. The WTs are given a value based on the MREF, and then Equation 3 is applied to the flask samples? If so, wouldn't the WTs by definition have to be stable? Or do you mean that they are stable relative to one another? Do the authors have a comment as to why 4845 is so variable?

L275: I think this is a little misleading, since changing MREF cylinders leads to large offsets in the record. I agree that generally based on Figure 3 the scale looks stable after the correction, but as I understand what the authors describe they are blind to WT drift. Or?

L290: But the primaries look systematically low (7002 and 7003), and 7008 shows clear drift. I would strike this sentence ("The ensemble thus suggests...")

L305: One primary is clearly drifting relative to the other two, does it really make sense to include this tank in the ensemble? Also, one would expect cylinders to drift over time. If possible these effects should be accounted for in an uncertainty analysis.

L350: It shows drift in your scale only if the GOLLUM cylinders are not drifting. It could be that 7008 is stable and the other two are not, or that all the cylinders are drifting together and 7008 is drifting slightly less or more than them. Without absolute constraints, it is unclear.

L404: "exact multitude of years" -- what does this mean?

L465: I think this seasonal cycle section (and section 4.1) is perhaps out of scope for the journal. Suggest to cut.

L486: "has been under much closer controlled thanks" -- should read e.g. "has been more closely controlled"

L492: I do not know what a valve cap is. Surely it is the o-ring which causes the seal? Why would permeation through the o-ring be impacted by an external cap?

L496: This is great to see, could you include some actual figures or numbers here?

L500: I find it hard to believe the trend could be impacted by sampling bias, particularly since Mace Head is sampled/filtered for background conditions.

L505-510: The decadal trend in APO should be virtually identical between two European background sites, and can't be explained by country-level differences in fossil fuel use. The authors acknowledge this on L509-510. Suggest this be cut.

L516: I agree that North Atlantic ventilation signals are likely to be present, but there are numerous other possible causes contributing to the different trends visible in APO between

the two sites. I think it would be better to simply state in a sentence or two that the authors do not know the cause of the difference between the two stations, and that the list of possible explanations include: changing continental fossil fuel signals, shifts in atmospheric transport, different degrees of sensitivity to North Atlantic ventilation, other possible causes, or an artifact(s) in the data. I think it's important to acknowledge the last one here. I also wonder how much of the curvature at MHD is due to the fit itself--maybe the annual averages aren't actually that different? On this last point, I think calculating the terrestrial and oceanic sinks is out of scope for this paper. I suggest cutting this section and the figures.

Section 5.2: I also think this is out of scope. The comparison of seasonal amplitudes between sites tells us really nothing about the quality of the data, since we expect there to be station to station differences. It also seems odd to me for an ESSD paper to briefly present data not pertaining to the dataset being published, such as the Weybourne or Halley measurements by other groups. Suggest cutting the entire section.

Figure 1: Three more panels showing the local site (e.g. satellite, street view, or topography) would be helpful here.

Figure 6 and 7: It is difficult to see the actual data because the fitted curves are on top of the points. I would suggest removing the curves completely and let the data speak for itself. I would also suggest zooming in on the CO<sub>2</sub> data a bit more.